Real Eigenvalue Extraction and the Distance to Uncontrollability

Emre Mengi
Computer Science Department
Courant Institute of Mathematical Sciences
New York University
mengi@cs.nyu.edu

May 22nd, 2006
Real Eigenvalue Problems

In control theory and numerical linear algebra various algorithms require real, imaginary or unit eigenvalues of large matrices.

- **Distance to uncontrollability of a pair** \((A, B)\) with \(A \in \mathbb{C}^{n \times n}\) and \(B \in \mathbb{C}^{n \times m}\):
  Gu’s method and its variants depend on the extraction of the real eigenvalues of problems of size \(O(n^2)\), i.e. they look for real \(\alpha\) such that

\[
\det(I \otimes F(\alpha) - F(\alpha + \eta)^T \otimes I) = 0
\]

where

\[
F(\alpha) = \begin{bmatrix}
\hat{B} & A - \alpha I \\
A^* - \alpha I & -\delta I
\end{bmatrix}
\]
• **Sep-λ of a pair** \((A, B)\) with \(A \in \mathbb{C}^{n \times n}\) and \(B \in \mathbb{C}^{p \times p}\):

Gu and Overton’s method seeks for real \(x\) such that

\[
\text{det}(I \otimes G_A(x) - G_B(x)^T \otimes I) = 0
\]

where

\[
G_A(x) = \begin{bmatrix}
A - xI & -\varepsilon I \\
\varepsilon I & -A^* + xI
\end{bmatrix}
\]


• **Hamiltonian eigenvalue problems:**

Hamiltonian matrices appear frequently in control theory. A Hamiltonian matrix \(H \in \mathbb{C}^{2n \times 2n}\) satisfying the property

\[
HJ = (HJ)^*, \quad J = \begin{bmatrix}
0 & I \\
-I & 0
\end{bmatrix}
\]

or a Hamiltonian pencil \(H_1 - \lambda H_2\) with the property

\[
H_1JH_2^T = H_2JH_1^T
\]

has eigenvalues in pairs \((\lambda, -\bar{\lambda})\). Often extraction of the imaginary eigenvalues of Hamiltonian problems is a crucial step for numerical algorithms.
Symplectic eigenvalue problems:

A symplectic pencil $S_1 - \lambda S_2$ with $S_1, S_2 \in \mathbb{C}^{2n \times 2n}$ satisfying

$$S_1JS_1^T = S_2JS_2^T$$

has eigenvalues in pairs $(\lambda, 1/\bar{\lambda})$. The unit eigenvalues of the symplectic problems are of interest in the applications. The symplectic pencil $S_1 - \lambda S_2$ can be mapped to a Hamiltonian pencil via a Cayley transformation.

$$H_1 - \lambda H_2 = (S_2 + S_1) - (S_2 - S_1)\lambda.$$
Overview

• A divide and conquer technique for real eigenvalue extraction
  – Algorithm description
  – Average case and worst case performance
  – Practical issues

• Distance to uncontrollability
  – Problem definition
  – A variant of Gu’s algorithm
  – Application of the real eigenvalue extraction
A Divide and Conquer Algorithm

- The algorithm is based on an efficient and accurate computation of the closest eigenvalue of $X \in \mathbb{C}^{p \times p}$ to a given point in the complex plane.

- The closest eigenvalue can be computed via
  - inverse Iteration or
  - a shift and invert preconditioned Arnoldi

if given a vector $v \in \mathbb{C}^p$, the multiplication

$$(X - \nu I)^{-1} v = u$$

can be performed efficiently.
A Divide and Conquer Algorithm

To extract the real eigenvalues of $\mathcal{X}$ contained in an interval $I = [l, u]$, 

- compute the eigenvalue $\lambda$ closest to the midpoint $\nu_m$ of $I$ (add $\lambda$ to the set of desired eigenvalues in case $\lambda$ is real),
- repeat the same procedure in the left interval $I_l = [l, \nu_m - |\lambda - \nu_m|]$ and in the right interval $I_r = [\nu_m + |\lambda - \nu_m|, u]$. 

Emre Mengi, Courant Institute
Initially call \texttt{Divide\_And\_Conquer} \((A, -D, D)\) where \(D\) is an upper bound on the norm of \(X\).

\texttt{Divide\_And\_Conquer} \((X, u, l)\)

[1.] Set the shift \(\nu_m \leftarrow \frac{(u+l)}{2}\).

[2.] Compute the eigenvalue \(\lambda\) closest to the shift \(\nu_m\).

\texttt{if} \(u-l < 2|\lambda - \nu_m|\) \texttt{then}

% Base case: there is no real eigenvalue in the interval \([l,u]\).

\texttt{Return \([\quad]\).}

\texttt{else}

% Recursive case: Search the left and right intervals.

\(\Lambda_l \leftarrow \texttt{Divide\_And\_Conquer}(X, l, \nu_m - |\lambda - \nu_m|)\)

\(\Lambda_r \leftarrow \texttt{Divide\_And\_Conquer}(X, \nu_m + |\lambda - \nu_m|, u)\)

\texttt{if} \(\lambda\) is real \texttt{then}

\texttt{Return} \(\lambda \cup \Lambda_l \cup \Lambda_r\).

\texttt{else}

\texttt{Return} \(\Lambda_l \cup \Lambda_r\).

\texttt{end if}

\texttt{end if}
A Divide and Conquer Algorithm

- Worst case: the algorithm can iterate at most $2p + 1$ times

- The progress of the algorithm can be represented by a full binary tree with at most $p + 1$ leaves.
- A full binary tree with $p + 1$ leaves has $p$ internal nodes.
- Each node of the tree corresponds to a closest eigenvalue computation.
A Divide and Conquer Algorithm

- **Average case:** Under the assumption that the eigenvalues are uniformly distributed, on average the number of closest eigenvalue computation is $O(\sqrt{p})$.

Let $E_j(X)$ denote the expected number of iterations given there are $j$ eigenvalues inside the circle of radius $D$.

**Theorem 1** (Average case for the divide and conquer algorithm). *Suppose the eigenvalues of the matrices input to the divide and conquer algorithm are selected uniformly and independently inside the circle of radius $\mu > D$. Then the expectation $E_j(X)$ is bounded above by $c\sqrt{j+d} - 1$ for all $c \geq \sqrt{12}$ and $d \in [4/c^2, 1/3]$. 
A Divide and Conquer Algorithm

Practical issues:

- We use ARPACK for the computation of the closest eigenvalue which is an implementation of the Arnoldi method.
  - There is no result stating that shift and invert preconditioned Arnoldi converges to the closest eigenvalue, but in practice this is almost always the case.
  - The convergence problems are observed especially when the norm of $X$ is big.

- Choosing $D$ very large does not alter the number of closest eigenvalue computations significantly. But this may cause convergence problems for ARPACK.
Distance to Uncontrollability

- Given a continuous first order dynamical system with the coefficient matrices $A \in \mathbb{C}^{n \times n}$ and $B \in \mathbb{C}^{n \times m}$

\[ x' = Ax + Bu, \quad x(0) = c_0 \]

where $x : \mathbb{R} \rightarrow \mathbb{C}^n$ and $u : \mathbb{R} \rightarrow \mathbb{C}^m$ are state and input functions of time respectively.

- We have full control over the system if the system is controllable, that is the system can be driven into any state at any given time $\tau$.

- Equivalent characterizations:

\[ \text{rank}[B \ AB \ A^2B \ldots A^{n-1}B] = n, \]

or

\[ \forall \lambda \in \mathbb{C} \quad \text{rank}[A - \lambda I \ B] = n. \]
Distance to Uncontrollability

- Norm of the smallest perturbation yielding an uncontrollable system (singular value characterization due to Eising)

\[
\tau(A, B) = \inf\{\|\Delta A \Delta B\| : (A + \Delta A, B + \Delta B) \text{ is uncontrollable}\} \\
= \inf_{\lambda \in \mathbb{C}} \sigma_n [A - \lambda I B].
\]
Distance to Uncontrollability

- Trisection algorithm by Burke, Lewis and Overton for arbitrary precision (variant of the bisection algorithm introduced by Gu).

1. Initialize $U = \sigma_n[A \ B]$ and $L = 0$.
2. If $U - L < \text{tol}$, terminate with $L$ and $U$.
3. Set $\delta_1 = L + 2/3(U - L)$ and $\delta_2 = L + 1/3(U - L)$.
4. If a $(\delta, \eta)$-pair exists, update the upper bound $U = \delta_1$. Otherwise update the lower bound $L = \delta_2$.
5. Go to step 2.
(\(\delta, \eta\))-pair: a pair of complex points \(z, z + \eta\) such that \(\sigma_{\min}[A - z I B] = \sigma_{\min}[A - (z + \eta) I B] = \delta\).
Distance to Uncontrollability

- **A \((\delta, \eta)\)-pair of \((A, B)\) exists:** Using the definition \(\tau(A, B) = \inf_{\lambda \in \mathbb{C}} \sigma_n [A - \lambda I \ B]\) we deduce
  \[
  \delta_1 = \delta \geq \tau(A, B).
  \]

- **No \((\delta, \eta)\)-pair of \((A, B)\) exists:** The result below based on the fact that singular values are well-conditioned implies \(\eta > 2(\delta - \tau(A, B))\) meaning
  \[
  \tau(A, B) > \delta - \eta/2 = \delta_2.
  \]
Theorem 2. Let $\delta \geq \tau$. For all $\xi \in [0, 2(\delta - \tau)]$, there exists a pair of real numbers $(\alpha, \beta)$ such that

\[
\sigma_n[A - (\alpha + \beta i) I \ B] = \delta, \quad \sigma_n[A - (\alpha + \xi + \beta i) I \ B] = \delta.
\]

$\delta$-pseudospectrum

$\eta_1, \eta_2 \geq \delta - \tau$
Distance to Uncontrollability

• Checking the existence of a \((\delta, \eta)\)-pair:

  – The rectangular matrix \([A - (\alpha + \beta i)I \ A] B\) has \(\delta\) as a singular value if and only if \(i\beta\) is an eigenvalue of the Hamiltonian matrix

\[
H(\alpha) = \begin{bmatrix}
-A^* + \alpha I & \delta I \\
BB^*/\delta - \delta I & A - \alpha I
\end{bmatrix}.
\]

  – If the points \(\alpha + \beta i\) and \(\alpha + \eta + \beta i\) is a \((\delta, \eta)\)-pair, the matrices \(H(\alpha)\) and \(H(\alpha + \eta)\) must share an imaginary eigenvalue namely \(\beta i\) meaning for some \(X \neq 0\)

\[
H(\alpha)X - XH(\alpha + \eta) = 0
\]

\[
(I \otimes H(\alpha) - H(\alpha + \eta)^T \otimes I)vec(X) = 0.
\]

  – This can be converted into a standard eigenvalue problem \(A_v = \alpha v\) of size \(2n^2\) whose real eigenvalues we seek.
Distance to Uncontrollability

- Checking the existence of a \((\delta, \eta)\)-pair:
  1. Extract the real eigenvalues of \(A\).
  2. For each real eigenvalue \(\alpha\), check whether \(H(\alpha)\) and \(H(\alpha + \eta)\) share an imaginary eigenvalue.
  3. If both 1 and 2 succeeds, the existence is verified. Otherwise it is failed.

- The complexity of the method is typically \(O(n^6)\) when the QR algorithm is used to compute the eigenvalues of \(A\).
The multiplication \((A - \nu I)^{-1}u = v\) can be performed efficiently

- Reversing the process from which the eigenvalue problem is obtained yields the Sylvester equation

\[
\begin{bmatrix}
A^* - \nu I & \delta I \\
BB^*/\delta - \delta I & -A + \nu I
\end{bmatrix}
\begin{bmatrix}
V_1 & W_1 \\
W_2 & V_2
\end{bmatrix}
+ \begin{bmatrix}
V_1 & W_1 \\
W_2 & V_2
\end{bmatrix}
\begin{bmatrix}
A - (\eta + \nu)I & BB^*/\delta - \delta I \\
\delta I & -A^* + (\eta + \nu)I
\end{bmatrix}
= 2
\begin{bmatrix}
U_1 & 0 \\
0 & -U_2
\end{bmatrix}.
\]

- The right-hand side is provided by the vector \(u = \begin{bmatrix}
vec(U_1) \\
vec(U_2)
\end{bmatrix} \in \mathbb{C}^{2n^2}\) that will be multiplied onto \((A - \nu I)^{-1}\)

- The product \(v = \begin{bmatrix}
vec(V_1) \\
vec(V_2)
\end{bmatrix} \in \mathbb{C}^{2n^2}\) can be obtained from the upper left and lower right blocks of the solution of the Sylvester equation.
Distance to Uncontrollability

- Under the assumption that each closest eigenvalue computation requires the matrix vector multiplication constant number of times, the cost of a closest eigenvalue computation is $O(n^3)$.

- When the real eigenvalue extraction technique is used,
  - the overall average running time reduces to $O(n^4)$,
  - the worst case running time reduces to $O(n^5)$. 
Distance to Uncontrollability

- Comparison of the running times on the normally distributed matrices:
  - Old method with QR to compute the eigenvalues of $A$.
    
    | (n,m) | (Lower bound, Upper bound) | Running time |
    |-------|-----------------------------|--------------|
    | (10,5) | (0.70316,0.70326] | 76           |
    | (25,10) | (0.92987,0.92994] | 24226        |
    | (30,23) | (1.99786,1.99794] | 62655        |

  - New method with shifted inverse iteration to compute the eigenvalues of $A$.
    
    | (n,m) | (Lower bound, Upper bound) | Running time | ave. calls to eigs |
    |-------|-----------------------------|--------------|-------------------|
    | (10,5) | (0.70316,0.70326] | 163          | 21.7              |
    | (25,10) | (0.92987,0.92994] | 1359         | 42.5              |
    | (30,23) | (1.99786,1.99794] | 3259         | 59.9              |
    | (40,10) | (0.83408,0.83418] | 6819         | 65.5              |
    | (45,40) | (2.79912,2.79920] | 8174         | 48.0              |
    | (50,13) | (1.07626,1.07632] | 15008        | 69.8              |
Distance to Uncontrollability

- Eigenvalue distributions of $\mathcal{A}$ for a normal matrix pair $(n = 5, m = 3)$ for various $\eta$

$\eta = 0.7914, \delta = 0.7914$

$\eta = 0.2345, \delta = 0.8940$
$\eta = 0.0695, \delta = 0.9765$

$\eta = 0.0137, \delta = 0.9593$
Distance to Uncontrollability

- Accuracy issues:
  - For more precision $\eta$ needs to be set smaller. As $\eta$ goes to zero, the norm of $\mathcal{A}$ approaches $\infty$.
  
  * Norm of $\mathcal{A}$ for a normally distributed pair and for various $\eta$

<table>
<thead>
<tr>
<th>$\eta$</th>
<th>0.7914</th>
<th>0.1563</th>
<th>0.0206</th>
<th>0.0027</th>
<th>0.0005</th>
</tr>
</thead>
<tbody>
<tr>
<td>norm of $\mathcal{A}$</td>
<td>118</td>
<td>439</td>
<td>3085</td>
<td>23510</td>
<td>118820</td>
</tr>
</tbody>
</table>

- Therefore the eigenvalues of $\mathcal{A}$ cannot be computed accurately in an absolute sense.

$$\delta \lambda = O(\kappa_{\lambda}(\mathcal{A}) \| \mathcal{A} \| \varepsilon_{\text{mach}})$$

- When the norm of $\mathcal{A}$ is large, inverse iteration and Arnoldi have convergence problems, since the matrix vector multiplications cannot be performed accurately.
Other applications

• The complexity of the algorithm by Gu and Overton for the sep-\(\lambda\) of a pair \(A \in \mathbb{C}^{n \times n}\) and \(B \in \mathbb{C}^{p \times p}\) reduces from \(O(n^7)\) to \(O(n^5)\) on average.

\[
sep_\lambda(A, B) = \inf \{ \max(\|\Delta A\|, \|\Delta B\|) : \Delta A \in \mathbb{C}^{n \times n}, \Delta B \in \mathbb{C}^{p \times p} \text{ s.t. } A + \Delta A \text{ and } B + \Delta B \text{ share an eigenvalue} \}
\]

• The computation of the distance to instability of a large and sparse matrix can be improved whenever the sparse \(LU\) decomposition can be performed efficiently.

\[
\gamma(A) = \inf \{ \Delta A : \Delta A \in \mathbb{C}^{n \times n} \text{ s.t. } (A + \Delta A) \text{ has an unstable eigenvalue} \}
\]